

# ENTROPY FOR SYMBOLIC DYNAMICS WITH OVERLAPPING ALPHABETS

FABIO DRUCKER, DAVID RICHESON, AND JIM WISEMAN

**ABSTRACT.** We consider shift spaces in which elements of the alphabet may overlap nontransitively. We define a notion of entropy for such spaces, give several techniques for computing lower bounds for it, and show that it is equal to a limit of entropies of (standard) full shifts. When a shift space with overlaps arises as a model for a discrete dynamical system with a finite set of overlapping neighborhoods, the entropy gives a lower bound for the topological entropy of the dynamical system.

## 1. INTRODUCTION

There is a long history of using symbolic dynamics to model more complicated dynamical systems. Suppose, for example, that we want to understand the dynamics of a continuous dynamical system  $f : X \rightarrow X$  on a compact metric space  $X$ . If we have compact sets  $N_1, \dots, N_n \subset X$  and an  $f$ -orbit  $(x, f(x), f^2(x), \dots)$  that remains in  $\bigcup_i N_i$ , then there is a symbol sequence  $(i_0, i_1, \dots) \in \{1, \dots, n\}^{\mathbb{N}}$  called an *itinerary* such that  $f^j(x) \in N_{i_j}$  for all  $j \geq 0$  (the itinerary may not be unique). We can then use the properties of the set of itineraries (a shift space) to tell us about the dynamics of  $f$ .

One example of this approach is a Markov partition, in which the sets  $N_1, \dots, N_n \subset X$  cover the entire space, overlap only on their boundaries, and map across each other in topologically simple ways ([6, §9.6]). In this case the resulting shift space of itineraries is a subshift of finite type. Another example is when the  $N_i$ 's are pairwise disjoint Conley index pairs, resulting in a cocyclic subshift ([8, 2, 3]).

While these approaches have been very fruitful, there are limitations with each of them. Markov partitions can be very difficult to construct in practice, and the Conley index pairs must be disjoint, making it impossible to cover a connected space.

In this article we allow the  $N_i$  to have nontrivial intersections. The complication that arises if we allow the  $N_i$ 's to intersect is that an orbit may have multiple itineraries. For example, a fixed point in  $N_1 \cap N_2$  could have itineraries  $(1, 1, 1, \dots)$ , or  $(1, 2, 1, 2, \dots)$ , or any sequence of 1's and 2's. Our task is to extract useful dynamical information about  $f$  from the shift space while being careful about the nonuniqueness of itineraries.

We begin with the definition of our object of study—a shift space with overlaps.

**Definition 1.** A *shift space with overlaps* is a pair  $(\Sigma, \mathcal{I})$  in which  $\Sigma$  is a one-sided shift space on the alphabet  $\{1, \dots, n\}$  and  $\mathcal{I}$  is a simple graph with vertex set  $\{1, \dots, n\}$ . We call  $\mathcal{I}$  an *intersection graph*. Two words  $(a_0, a_1, \dots), (b_0, b_1, \dots) \in \Sigma$  (finite or infinite) are *indistinguishable* iff there is an  $\mathcal{I}$ -edge between the vertices  $a_i$  and  $b_i$  for all  $i$ .

Shift spaces with overlaps are interesting objects to study on their own, but most often they are a model for an existing dynamical system. Thus we have the following definition.

**Definition 2.** A compact metric space  $X$ , a continuous map  $f : X \rightarrow X$ , and a collection of nonempty compact sets  $N_1, \dots, N_n \subset X$  is called a *dynamical realization* of  $(\Sigma, \mathcal{I})$  provided:

- (1) If  $N_i \cap N_j \neq \emptyset$  and  $i \neq j$ , then there is an edge from  $i$  to  $j$  in  $\mathcal{I}$ .
- (2) If  $(a_0, a_1, \dots)$  is a (finite or infinite) word in  $\Sigma$ , then there is a point  $x \in X$  such that  $f^i(x) \in \text{Int}(N_{a_i})$  for all  $i$ .

In practice, we use topological methods to verify that (2) holds. The methods may be elementary, such as boxes that stretch across other boxes, as in the case of Markov partitions or their topological generalizations shown in Figure 1. Or they may be more sophisticated; for example, in [5] we use techniques from Conley index theory to verify the property.

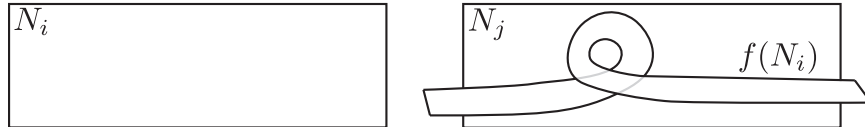


FIGURE 1.

For much of this article we assume that  $\Sigma$  is a subshift of finite type. When  $\Sigma$  is the vertex shift associated to a transition graph  $\mathcal{T}$  we superimpose the graphs  $\mathcal{T}$  and  $\mathcal{I}$  with the edges of  $\mathcal{T}$  being solid arrows and the edges of  $\mathcal{I}$  dashed segments (see Figure 2, for example); we will refer to this merged graph as  $\mathcal{TI}$ .

We are using  $(\Sigma, \mathcal{I})$  as an abstract model of our dynamical system  $f : X \rightarrow X$ . As such, two indistinguishable elements in  $\Sigma$  could be itineraries of the same point in  $X$ . Thus in a sense we must treat two indistinguishable elements in  $\Sigma$  as the same point. However, what makes this scenario interesting is that indistinguishability is not a transitive property, and is thus not an equivalence relation. In general,  $(\Sigma, \mathcal{I})$  is not a topological space. (It is a tolerance space—see, for example, [7].)

Our goal is to extract information about the dynamical complexity of  $f$  from  $(\Sigma, \mathcal{I})$ , and most often in this article, from the graph  $\mathcal{TI}$ . We define a notion of topological entropy for shifts with overlap,  $h_{\mathcal{I}}(\Sigma)$ , in Section 2 and show that it is a lower bound for the entropy of a dynamical realization  $f$ .

We estimate  $h_{\mathcal{I}}(\Sigma)$  using only disjoint sets in Section 3, discuss decompositions of the shift space in Section 4, and estimate  $h_{\mathcal{I}}(\Sigma)$  using connected components in Section 5. In Section 6 we discuss higher shifts, and in Section 7 we prove that  $h_{\mathcal{I}}(\Sigma)$  can be computed as a limit of the entropies of ordinary (nonoverlapping) embedded shifts.

One of the motivations for undertaking this study is its potential application to computational dynamical systems. There is an active research program of using the Conley index and computational topology to give rigorous entropy bounds for dynamical systems. In particular, one could use rigorous computation to obtain an index system (as described in [5]) and from that a shift space with overlaps. So it is important to find techniques for computing the entropy, or bounds on the entropy, of shift spaces with overlaps.

## 2. TOPOLOGICAL ENTROPY OF SHIFT SPACES WITH OVERLAP

We are primarily interested in using  $(\Sigma, \mathcal{I})$  to help us obtain a lower bound for the entropy of  $f$ . Despite the fact that  $(\Sigma, \mathcal{I})$  is not a topological space, we can make a reasonable definition of its topological entropy, and we prove that it is a lower bound for the entropy of  $f$  (Theorem 3).

If we ignore the intersections, then  $\Sigma$  is a shift space and we can compute  $h(\Sigma)$  by computing the growth rate of the number of words of length  $n$  as  $n$  goes to infinity ([4, §4.1]). Let  $\mathcal{B}_n(\Sigma)$  be the set of words of length  $n$ ; then the topological entropy of  $\Sigma$  is

$$h(\Sigma) = \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n(\Sigma)|.$$

Recall that if the shift space  $\Sigma$  is a subshift of finite type with transition graph  $\mathcal{T}$ , then there is a particularly easy way to compute the entropy. The graph  $\mathcal{T}$  has an associated *transition matrix*,  $A_{\mathcal{T}}$ , in which the  $(i, j)$ -entry is 1 if there is an edge from the  $i$ th vertex to the  $j$ th vertex and is 0 otherwise. Then

$$h(\Sigma) = \log(\lambda),$$

where  $\lambda$  is the Perron eigenvalue of  $A_{\mathcal{T}}$ —the unique largest real eigenvalue of  $A_{\mathcal{T}}$ .

In our setting, we want to count the number of words of length  $n$  that are mutually distinguishable, which in general is less than  $|\mathcal{B}_n(\Sigma)|$ . (For example, if vertices 1 and 2 overlap, then we would not be able to distinguish the words  $(1, 3)$  and  $(2, 3)$ .) Let  $B$  be a set of words in  $\Sigma$  of length  $n$ . We say that  $B$  is *n-separated* if no two words in  $B$  are indistinguishable. Let  $\mathcal{B}_n^{\text{sep}}(\Sigma, \mathcal{I})$  be an  $n$ -separated set of maximum cardinality. We define  $h_{\mathcal{I}}(\Sigma)$ , the *topological entropy of a shift space with overlaps*, to be

$$h_{\mathcal{I}}(\Sigma) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n^{\text{sep}}(\Sigma, \mathcal{I})|.$$

The following theorem provides the motivation for studying the entropy of shifts with overlaps.

**Theorem 3.** *If  $f$  is a dynamical realization of  $(\Sigma, \mathcal{I})$ , then  $h(f) \geq h_{\mathcal{I}}(\Sigma)$ .*

*Proof.* Let  $\varepsilon = \min d(N_i, N_j)$ , where the minimum is over all disjoint pairs  $N_i$  and  $N_j$ . Then an  $n$ -separated word in  $\Sigma$  corresponds to an  $(n, \varepsilon)$ -separated orbit for  $f$ . Since  $h(f) = \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log r(n, \varepsilon)$ , where  $r(n, \varepsilon)$  is the maximum cardinality of a set of  $(n, \varepsilon)$ -separated orbits ([6, §8.1]), the result follows.  $\square$

The rest of the paper is devoted to computing or finding lower bounds for  $h_{\mathcal{I}}(\Sigma)$ .

### 3. INDEPENDENT SUBSETS OF THE INTERSECTION GRAPH

In this section we take advantage of the fact that if all words are distinguishable, then we can treat our shift space with overlaps as an ordinary shift space.

**Proposition 4.** *Let  $(\Sigma, \mathcal{I})$  be a shift space with overlaps in which the edge set of  $\mathcal{I}$  is empty. Then  $h_{\mathcal{I}}(\Sigma) = h(\Sigma)$ .*

**Corollary 5.** *Let  $(\Sigma, \mathcal{I})$  be a shift space with overlaps. If  $\Sigma' \subset \Sigma$  is a shift space in which all elements are pair-wise distinguishable, then  $h_{\mathcal{I}}(\Sigma) \geq h_{\mathcal{I}}(\Sigma') = h(\Sigma')$ .*

Let  $V(\mathcal{G})$  and  $E(\mathcal{G})$  denote the vertex and edge sets of a graph  $\mathcal{G}$ , respectively. Given a (nondirected) graph  $\mathcal{G}$ , a subset  $V \subset V(\mathcal{G})$  is *independent* if there are no edges between any pair of vertices in  $V$ . The maximum cardinality of an independent subset of  $\mathcal{G}$  is the *independence number* of  $\mathcal{G}$  and is denoted  $\text{ind}(\mathcal{G})$ .

For any  $V \subset V(\mathcal{I})$ , let  $\Sigma_V = \{(a_0, a_1, \dots) \in \Sigma : a_i \in V\}$ . Notice that if  $\Sigma$  is a shift space, then  $\Sigma_V$  is also a shift space. A *vertex-induced subgraph* of a graph  $\mathcal{G}$  is a subgraph  $\mathcal{G}'$  with the property that any  $\mathcal{G}$ -edge whose endpoints are in  $V(\mathcal{G}')$  is in  $E(\mathcal{G}')$ . If  $V \subset V(\mathcal{I})$ , let  $\mathcal{I}_V$  denote the vertex-induced subgraph of  $\mathcal{I}$  with vertex set  $V$ .

Corollaries 6 and 7 give us a tactic for finding a lower bound for  $h_{\mathcal{I}}(\Sigma)$ . Remove enough vertices of  $\mathcal{I}$  so that we have an independent set of vertices; then the corresponding shift space with overlaps has no overlaps—it is just a shift space. The entropy of this shift space is a lower bound for  $h_{\mathcal{I}}(\Sigma)$ .

**Corollary 6.** *If  $V \subset V(\mathcal{I})$  is an independent set, then  $h_{\mathcal{I}}(\Sigma) \geq h_{\mathcal{I}_V}(\Sigma_V) = h(\Sigma_V)$ .*

In the following corollaries we assume that  $\Sigma$  is a shift of finite type with transition graph  $\mathcal{T}$  and intersection graph  $\mathcal{I}$ .

**Corollary 7.** *Let  $V \subset V(\mathcal{I})$  be an independent set and let  $\mathcal{T}'\mathcal{I}'$  be the vertex-induced subgraph of  $\mathcal{T}\mathcal{I}$  with vertex set  $V$ . Then  $h_{\mathcal{I}}(\Sigma) \geq \log(\lambda)$  where  $\lambda$  is the Perron eigenvalue of  $A_{\mathcal{T}'}$ .*

**Corollary 8.** *Suppose  $\mathcal{T}$  is a complete digraph. Then  $h_{\mathcal{I}}(\Sigma) \geq \log(\text{ind}(\mathcal{I}))$ .*

*Proof.* Let  $V$  be a maximum cardinality independent subset of  $\mathcal{I}$ . The vertex-induced subgraph of  $\mathcal{T}$  with vertex set  $V$  is a complete digraph with  $\text{ind}(\mathcal{I})$  vertices. By Corollary 7,  $h_{\mathcal{I}}(\Sigma) \geq \log(\text{ind}(\mathcal{I}))$ .  $\square$

#### 4. DECOMPOSITION INTO IRREDUCIBLE AND PRIMITIVE COMPONENTS

We begin this section with some basic definitions from graph theory and linear algebra. (See [4, ch. 4] or any non-introductory linear algebra text for more details.)

Let  $\mathcal{G}$  be a directed graph. A *vertex path*  $(a_0, a_1, \dots, a_{l-1})$  of length  $l$  is a sequence of vertices in  $\mathcal{G}$  such that there is an edge from vertex  $a_i$  to  $a_{i+1}$  for each  $i$ ,  $0 \leq i \leq l-2$ . In what follows, we assume that  $\mathcal{G}$  has  $n$  vertices and no parallel edges and that every vertex has at least one edge leaving it and one edge entering it. Then the adjacency matrix for  $\mathcal{G}$ ,  $A_{\mathcal{G}}$ , is an  $n \times n$   $(0,1)$ -matrix with no row or column containing all zeros. Notice that for  $k > 1$  the matrix  $A_{\mathcal{G}}^k$  need not be a  $(0,1)$ -matrix. However, if the  $(i, j)$ -entry of  $A_{\mathcal{G}}^k$  is nonzero, then there exists a vertex path of length  $k+1$  from vertex  $i$  to vertex  $j$  (actually, the value of the  $(i, j)$  gives the number of such vertex paths).

A graph is *irreducible* provided there is a vertex path between any two vertices, and a matrix is *irreducible* if it is the adjacency matrix for an irreducible graph. It is always possible to decompose a graph into its irreducible components; the corresponding matrix is (after rearranging the order of the vertices, or equivalently, conjugating by a permutation matrix) in block-triangular form with each of the blocks on the diagonal the adjacency matrix for one of the irreducible components of the graph.

An irreducible graph  $\mathcal{G}$  may exhibit cyclic behavior. In particular for each vertex  $i$ ,  $\text{Per}(i) = \gcd\{n : \text{there is a vertex path of length } n+1 \text{ from } i \text{ to itself}\}$  exists and is the same value for all  $i$ . This common value is called the *period* of  $\mathcal{G}$ , denoted  $\text{Per}(\mathcal{G})$ . Define the *period* of an irreducible  $(0,1)$ -matrix  $A$ ,  $\text{Per}(A)$ , to be the period of the associated graph. If the period of the graph or the matrix is 1, then we call it *primitive*. If  $A$  is an irreducible matrix with  $p = \text{Per}(A)$ , then (after reordering vertices)  $A^p$  is a block triangular matrix with primitive matrices along the diagonal. Finally, if  $A$  is a primitive matrix, then  $A^k$  is eventually positive. The *index of primitivity* of  $A$ ,  $\gamma(A)$ , is the least integer  $k$  such that  $A^k > 0$ .

Now suppose our shift space with overlaps has an associated graph  $\mathcal{T}\mathcal{I}$ ; that is, we consider  $(\Sigma, \mathcal{I})$  where  $\Sigma$  is a subshift of finite type with transition graph  $\mathcal{T}$ . We make the standing assumption that every vertex of  $\mathcal{T}$  has at least one edge leaving it and one edge entering it; if not, we can remove the “stranded” vertices without affecting the dynamics. Recall that the graph-theoretical notions of irreducibility and primitivity have dynamical interpretations for the shift:  $\mathcal{T}$  is irreducible iff  $\Sigma$  is *topologically transitive*, and  $\mathcal{T}$  is primitive iff  $\Sigma$  is *topologically mixing*.

**Theorem 9.** *Suppose  $\mathcal{T}$  is primitive with index of primitivity  $\gamma$ . Then  $h_{\mathcal{I}}(\Sigma) \geq \frac{\log(\text{ind}(\mathcal{I}))}{\gamma}$ .*

*Proof.* Let  $\{b_1, \dots, b_{\text{ind}(\mathcal{I})}\}$  be an independent set of vertices. Since  $A^\gamma > 0$ , for any pair  $b_i, b_j$ , there is a word  $(b_i, a_2, a_3, \dots, a_\gamma, b_j)$  in  $\Sigma$ . More generally, for any  $n$  and any  $b_{i_1}, \dots, b_{i_n}$ , there is a word

$$(b_{i_1}, a_2, \dots, a_\gamma, b_{i_2}, a_{\gamma+2}, \dots, a_{2\gamma}, \dots, b_{i_n}, a_{(n-1)\gamma+2}, \dots, a_{n\gamma})$$

in  $\Sigma$ . The collection of all  $\text{ind}(\mathcal{I})^n$  such words is  $n\gamma$ -separated, so we have that  $|\mathcal{B}_{n\gamma}^{\text{sep}}(\Sigma, \mathcal{I})| \geq \text{ind}(\mathcal{I})^n$ , and thus

$$h_{\mathcal{I}}(\Sigma) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n^{\text{sep}}(\Sigma, \mathcal{I})| \geq \limsup_{n \rightarrow \infty} \frac{1}{n\gamma} \log |\mathcal{B}_{n\gamma}^{\text{sep}}(\Sigma, \mathcal{I})| \geq \frac{\log(\text{ind}(\mathcal{I}))}{\gamma}.$$

□

**Corollary 10.** *If  $\mathcal{T}$  is primitive with  $n$  vertices, then  $h_{\mathcal{I}}(\Sigma) \geq \frac{\log(\text{ind}(\mathcal{I}))}{n^2 - 2n + 2}$ .*

*Proof.* This follows from the fact that  $\gamma(\mathcal{T}) \leq n^2 - 2n + 2$  ([1, Thm. 4.14]).

□

In general  $\mathcal{T}$  is not primitive. In this case we break down  $\mathcal{T}$  into its irreducible components. Each irreducible component has some period  $p$ , and can thus be decomposed into  $p$  primitive components. We then apply Theorem 9 to obtain the following corollary.

**Corollary 11.** *Suppose  $\mathcal{T}' \subset \mathcal{T}$  is a primitive component with index of primitivity  $\gamma$  that resides in an irreducible component of period  $p$  and that  $\mathcal{I}'$  is the associated intersection graph. Then*

$$h_{\mathcal{I}}(\Sigma) \geq \frac{\log(\text{ind}(\mathcal{I}'))}{p\gamma}.$$

**Corollary 12.** *If there are two vertices of  $\mathcal{T}$  that are in the same primitive component and are not joined by an  $\mathcal{I}$ -edge, then  $(\Sigma, \mathcal{I})$  has positive entropy.*

**Example 13.** Let  $U \subset \mathbb{R}^2$  be an open set and  $f : U \rightarrow \mathbb{R}^2$  be a continuous function that maps the sets  $N_1, \dots, N_{11} \subset U$  as shown in Figure 2. The associated graph  $\mathcal{T}\mathcal{I}$  is shown on the right.

First we remove vertices 6 and 9 since they lack an outgoing and an incoming edge in  $\mathcal{T}$ , respectively. (Note: if we also removed two of the vertices 1, 2, and 4 to obtain an independent set, then we could apply Corollary 7, but the corresponding shift space would have zero entropy.) The resulting transition graph has two irreducible components, with vertex sets  $\{8\}$  and  $\{1, 2, 3, 4, 5, 7, 9, 10, 11\}$ . The second of these has period 2 and it decomposes into the primitive components  $\{3, 4, 7, 11\}$  and  $\{1, 2, 5, 10\}$ . It is straightforward to show that both of these have index of primitivity 4 and that the first of these is an independent set. Thus we conclude that  $h(f) \geq h_{\mathcal{I}}(\Sigma) \geq \frac{1}{2.4} \log(4) \approx 0.173$ . We will return to this example later.

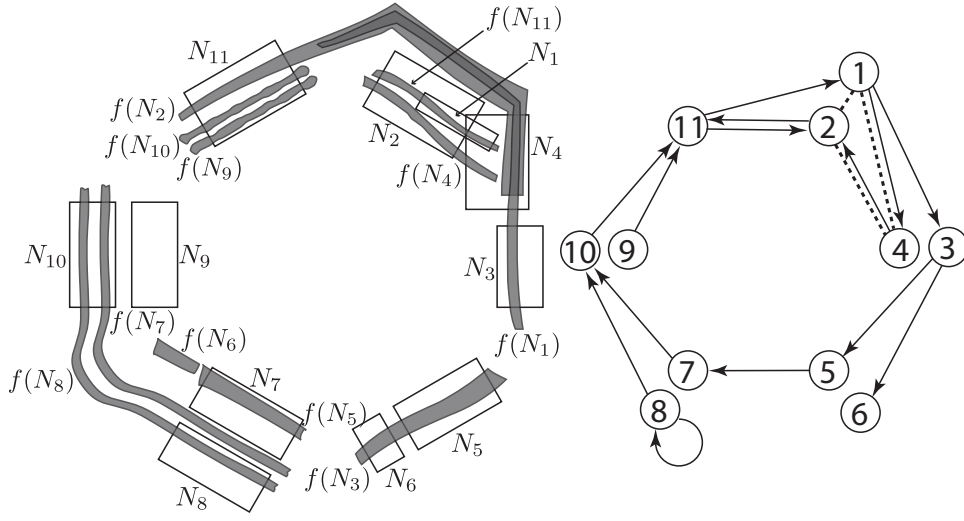


FIGURE 2.

**Example 14.** Let  $S^1$  be the circle, viewed as  $\mathbb{R}/\mathbb{Z}$  and let  $f : S^1 \rightarrow S^1$  be a map that is  $C^0$ -close to the doubling map ( $x \mapsto 2x$ ). Let  $N_1 = [-0.1, 0.35]$ ,  $N_2 = [0.15, 0.6]$ ,  $N_3 = [0.4, 0.85]$ , and  $N_4 = [0.65, 1.1]$ . The transition matrix and the graph  $\mathcal{TI}$  are shown in Figure 3.

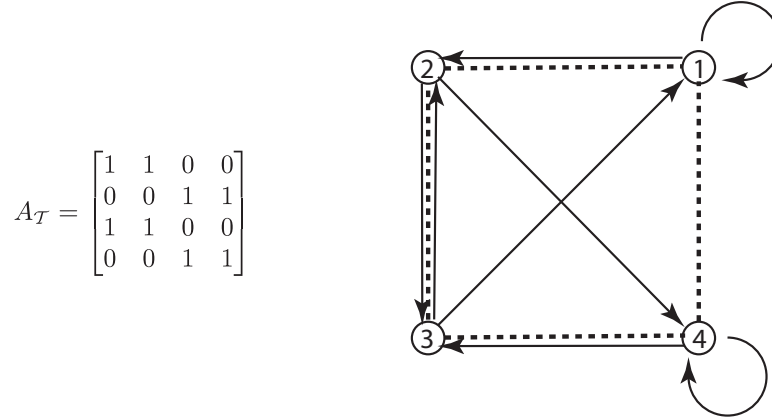


FIGURE 3.

Since  $A_{\mathcal{T}}^2$  consists of all 1's,  $A_{\mathcal{T}}$  is irreducible with  $\text{Per}(A_{\mathcal{T}}) = 1$  and primitive with  $\gamma(A_{\mathcal{T}}) = 2$ . The maximal independent sets in  $\mathcal{I}$  are  $\{1, 3\}$  and  $\{2, 4\}$ , so  $\text{ind}(\mathcal{I}) = 2$ . Thus  $h(f) \geq h_{\mathcal{I}}(\Sigma) \geq \frac{1}{2} \log(2)$ , which is half the entropy of the doubling map. We will return to this example.

5. SOFIC SHIFTS AND  $\mathcal{I}$ -COMPONENT SHIFTS

Vertices that are adjacent in  $\mathcal{I}$  are indistinguishable. This is a problem for our symbolic dynamics, since different symbols may not correspond to different points. One way of working around this problem is to give neighboring vertices the same label. In other words, give each vertex in a connected component of  $\mathcal{I}$  the same label.

Suppose  $\mathcal{I}$  has connected components  $A_1, \dots, A_r$ . Then create the shift space  $\Sigma_{\mathcal{I}}$  on the alphabet  $\{A_1, \dots, A_r\}$  as follows. The element  $(A_{i_0}, A_{i_1}, \dots) \in \Sigma_{\mathcal{I}}$  if and only if there is an element  $(a_0, a_1, \dots) \in \Sigma$  such that  $a_k \in A_{i_k}$  for all  $k$ . We call  $\Sigma_{\mathcal{I}}$  the  *$\mathcal{I}$ -component shift space* for  $(\Sigma, \mathcal{I})$ .

**Theorem 15.** *Let  $\Sigma_{\mathcal{I}}$  be the  $\mathcal{I}$ -component shift space associated to  $(\Sigma, \mathcal{I})$ . Then  $h_{\mathcal{I}}(\Sigma) \geq h(\Sigma_{\mathcal{I}})$ .*

*Proof.* Let  $\mathcal{B}_m(\Sigma_{\mathcal{I}})$  be the set of words of length  $m$  in  $\Sigma_{\mathcal{I}}$ . For each element  $(A_{i_0}, \dots, A_{i_{m-1}}) \in \mathcal{B}_m(\Sigma_{\mathcal{I}})$ , pick one word  $(a_0, \dots, a_{m-1})$  in  $\Sigma$  such that  $a_k \in A_{i_k}$  for  $k = 0, \dots, m-1$ . The collection of these words in  $\Sigma$  are  $m$ -separated. However, it may not be a maximal  $m$ -separated set. Thus  $|\mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I})| \geq |\mathcal{B}_m(\Sigma_{\mathcal{I}})|$ , and hence  $h_{\mathcal{I}}(\Sigma) \geq h(\Sigma_{\mathcal{I}})$ .  $\square$

A set of vertices in a nondirected graph form a *clique* if every pair of vertices in the set are joined by an edge.

**Proposition 16.** *If the vertices in each connected component of  $\mathcal{I}$  form a clique, then  $h_{\mathcal{I}}(\Sigma) = h(\Sigma_{\mathcal{I}})$ .*

*Proof.* By Theorem 15, all we must prove is that  $h_{\mathcal{I}}(\Sigma) \leq h(\Sigma_{\mathcal{I}})$ . Let  $\mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I})$  be a maximal  $m$ -separated set. Consider the function  $\psi : \mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I}) \rightarrow \mathcal{B}_m(\Sigma_{\mathcal{I}})$  given by  $\psi(a_0, a_1, \dots, a_{m-1}) = (A_{i_0}, A_{i_1}, \dots, A_{i_{m-1}})$  where  $a_k \in A_{i_k}$  for all  $k$ . We will prove that this function is injective. Suppose  $\psi(a_0, a_1, \dots, a_{m-1}) = \psi(b_0, b_1, \dots, b_{m-1})$  for some  $(a_0, a_1, \dots, a_{m-1}), (b_0, b_1, \dots, b_{m-1}) \in \mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I})$ . Then  $a_k, b_k \in A_{i_k}$  for all  $k$ . But the subgraphs of  $\mathcal{I}$  with vertex sets  $A_{i_k}$  are cliques, so  $a_k$  and  $b_k$  are the same or are indistinguishable. Thus  $(a_0, a_1, \dots, a_{m-1})$  and  $(b_0, b_1, \dots, b_{m-1})$  are indistinguishable. Since  $\mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I})$  is  $m$ -separated,  $(a_0, a_1, \dots, a_{m-1}) = (b_0, b_1, \dots, b_{m-1})$ , and  $\psi$  is injective. Because  $\psi$  is an injective function between finite sets, we conclude that  $|\mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I})| \leq |\mathcal{B}_m(\Sigma_{\mathcal{I}})|$ . Hence  $h_{\mathcal{I}}(\Sigma) \leq h(\Sigma_{\mathcal{I}})$ .  $\square$

An important special case occurs when  $\Sigma$  is a subshift of finite type. Then  $\Sigma_{\mathcal{I}}$  is a sofic shift (see [4, Ch. 3] for more about sofic shifts and the terms in this paragraph). If  $(\Sigma, \mathcal{I})$  is represented by the graph  $\mathcal{T}\mathcal{I}$ , then  $\Sigma_{\mathcal{I}}$  is the *sofic shift* represented by the directed graph obtained from  $\mathcal{T}$  by giving all vertices in a connected component of  $\mathcal{I}$  the same label. If this graph is *right-resolving* (that is, if every edge from a given vertex points to a different label), then the entropy of the sofic shift is simply the log of the Perron eigenvalue of the adjacency matrix for  $\mathcal{T}$ . If the graph is not right-resolving then there is a mechanical procedure for transforming it into



a right-resolving graph representing the same sofic shift. In particular, the entropy is easy to compute.

**Example 17.** Consider the graph  $\mathcal{TI}$  shown on the left in Figure 4. If we give vertices 1 and 3 the same label we obtain the “golden mean” shift shown on the right (a shift that is well-known to be sofic but not of finite type). By Proposition 16 the entropies of the two are equal, and since the graph for the sofic shift is right-resolving, we can easily compute the entropy from the adjacency matrix:  $h_{\mathcal{I}}(\Sigma) = h(\Sigma_{\mathcal{I}}) = \log(\frac{1}{2}(1 + \sqrt{5}))$ .

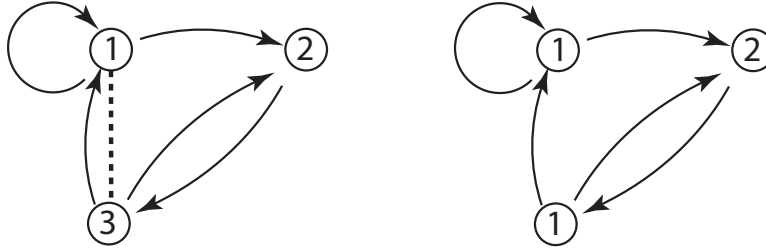


FIGURE 4.

**Example 18.** Consider the sofic shift associated to the shift space with overlaps in Example 13 (with vertices 6, 8, and 9 removed). That is, we give vertices 1, 2, and 4 the same label (the graph on the left in Figure 5). The graph is not right-resolving, but after being put in right-resolving form (the graph on the right) we find that  $h_{\mathcal{I}}(\Sigma) \geq \log(\frac{1}{6}(108 + 12\sqrt{69})^{1/3} + 2(108 + 12\sqrt{69})^{-1/3}) \approx .281$ , which is a larger lower bound than the one obtained in Example 13.

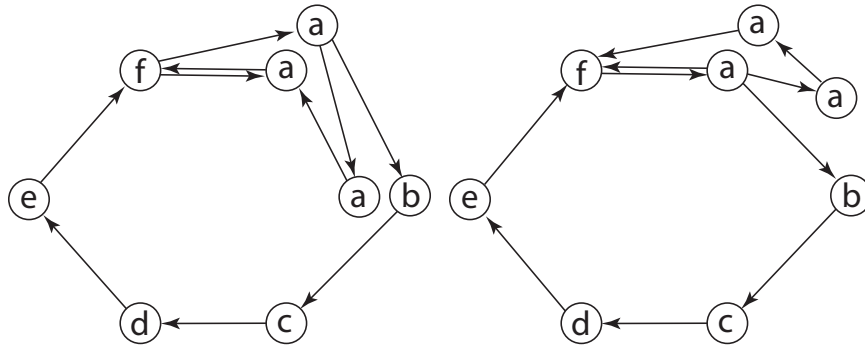


FIGURE 5.

**Example 19.** On the other hand, the sofic shift corresponding to the doubling map of Example 14 has only one symbol, so we obtain a lower bound of zero for the entropy.

In some circumstances we could obtain a larger lower bound for the entropy by using a combination of removing vertices (as described in Section 3) and relabeling (as described in this section) than we could using either of these techniques alone. We could remove a few vertices to increase the number of connected components of  $\mathcal{I}$ , then relabel to obtain an  $\mathcal{I}$ -component shift with high entropy.

## 6. HIGHER SHIFTS

In this section we examine higher shifts. This is a well-known way of altering the transition graph for a subshift of finite type to obtain a new graph which generates a topologically conjugate subshift of finite type. The higher vertex shift allows us to subdivide the set of vertices, thus presumably giving us a greater number of disjoint vertices. We show how to implement this notion for shift spaces with overlap generated from a graph  $\mathcal{TI}$ . We refer the reader to [4, Sect. 2.3] for details on higher shifts for edge-labeled graphs; the constructions can be modified easily for vertex-labeled graphs, which is what we need.

For  $m \geq 1$  we define the  $m$ th higher vertex graph  $\mathcal{T}_{[m]}$  to have vertex set equal to the collection of all vertex paths of length  $m$  in  $\mathcal{T}$ , with an edge from a vertex  $(i_0, \dots, i_{m-1})$  to a vertex  $(i_1, \dots, i_{m-1}, i_m)$  provided there is an edge in  $\mathcal{T}$  from  $i_{m-1}$  to  $i_m$ . Note that the first higher vertex shift  $\mathcal{T}_{[1]}$  is simply  $\mathcal{T}$ .

The intersection graph  $\mathcal{I}$  induces an intersection graph  $\mathcal{I}_{[m]}$  for the vertices of  $\mathcal{T}_{[m]}$ . There is an (undirected) edge between the vertices  $(i_0, \dots, i_{m-1})$  and  $(j_0, \dots, j_{m-1})$  if the words are indistinguishable; that is, if there is an edge in  $\mathcal{I}$  between  $i_k$  and  $j_k$  for all  $k$ .

**Theorem 20.** *If  $\mathcal{T}_{[m]}\mathcal{I}_{[m]}$  is the graph associated to the  $m$ th higher shift of  $(\Sigma_{\mathcal{T}}, \mathcal{I})$ , then  $h_{\mathcal{I}_{[m]}}(\Sigma_{\mathcal{T}_{[m]}}) = h_{\mathcal{I}}(\Sigma_{\mathcal{T}})$ .*

*Proof.* Let  $m \geq 1$  be fixed. For each  $n$ , there exists a natural bijective function  $\psi_n : \mathcal{B}_{n+m-1}(\Sigma_{\mathcal{T}}) \rightarrow \mathcal{B}_n(\Sigma_{\mathcal{T}_{[m]}})$ . Specifically,  $\psi_n(b_0, \dots, b_{n+m-2}) = (w_0, \dots, w_{n-1})$ , where  $w_i = (b_i, \dots, b_{m+i-1})$ . Moreover,  $\psi_n(b_0, \dots, b_{n-1})$  and  $\psi_n(c_0, \dots, c_{n-1})$  are distinguishable if and only if  $(b_0, \dots, b_{n-1})$  and  $(c_0, \dots, c_{n-1})$  are distinguishable. Thus,  $|\mathcal{B}_{n+m-1}^{\text{sep}}(\Sigma_{\mathcal{T}}, \mathcal{I})| = |\mathcal{B}_n^{\text{sep}}(\Sigma_{\mathcal{T}_{[m]}}, \mathcal{I}_{[m]})|$ , and hence

$$\begin{aligned} h_{\mathcal{I}_{[m]}}(\Sigma_{\mathcal{T}_{[m]}}) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_n^{\text{sep}}(\Sigma_{\mathcal{T}_{[m]}}, \mathcal{I}_{[m]})| \\ &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{B}_{n+m-1}^{\text{sep}}(\Sigma_{\mathcal{T}}, \mathcal{I})| \\ &= h_{\mathcal{I}}(\Sigma_{\mathcal{T}}). \end{aligned}$$

□

**Proposition 21.**  $\text{ind}(\mathcal{I}_{[m]}) \geq \text{ind}(\mathcal{I})$ .

*Proof.* (Recall that we are assuming that each vertex has at least one edge leaving it.) If  $\{b_1, \dots, b_{\text{ind}(\mathcal{I})}\}$  is an independent set of vertices of  $\mathcal{I}$ , and  $w_i = (b_i, a_{i_1}, \dots, a_{i_{m-1}})$  is, for each  $b_i$ , any word of length  $m$ , then  $\{w_1, \dots, w_{\text{ind}(\mathcal{I})}\}$  is an independent set of vertices of  $\mathcal{I}_{[m]}$ .  $\square$

**Proposition 22.** *Assume that  $\mathcal{T}$  has at least two vertices. If  $\mathcal{T}$  is primitive, then so is  $\mathcal{T}_{[m]}$ , and  $\gamma(\mathcal{T}_{[m]}) = \gamma(\mathcal{T}) - 1 + m$ . (If  $\mathcal{T}$  has only one vertex, then  $\mathcal{T}_{[m]}$  is isomorphic to  $\mathcal{T}$ .)*

*Proof.* A graph is primitive with exponent  $\gamma$  if and only if from any vertex to any other there is a vertex path of length  $\gamma + 1$ , and  $\gamma$  is the smallest number with this property. Observe that  $(i_{m-1}, k_1, \dots, k_p, j_0)$  is a vertex path in  $\mathcal{T}$  (of length  $p + 2$ , with  $p \geq 0$ ) if and only if

$$((i_0, \dots, i_{m-1}), (i_1, \dots, i_{m-1}, k_1), \dots, (k_p, j_0, \dots, j_{m-2}), (j_0, \dots, j_{m-1}))$$

is a vertex path in  $\mathcal{T}_{[m]}$  (of length  $p + m + 1$ ) for any vertices  $(i_0, \dots, i_{m-1})$  and  $(j_0, \dots, j_{m-1})$ . Between any two vertices of  $\mathcal{T}$  there is a vertex path of length  $p + 2 = \gamma(\mathcal{T})$ , so between any two vertices of  $\mathcal{T}_{[m]}$  there is a vertex path of length  $p + m + 1 = \gamma(\mathcal{T}) - 1 + m$ . Thus  $\gamma(\mathcal{T}_{[m]}) \leq \gamma(\mathcal{T}) - 1 + m$ .

To prove the opposite inequality, first assume that  $\gamma(\mathcal{T}) > 1$ . Then there exist two vertices of  $\mathcal{T}$  such that there is no vertex path from the first to the second of length  $\gamma(\mathcal{T})$ , and thus two vertices of  $\mathcal{T}_{[m]}$  with no vertex path of length  $\gamma(\mathcal{T}) - 1 + m$  between them, so  $\gamma(\mathcal{T}_{[m]}) \geq \gamma(\mathcal{T}) - 1 + m$ .

Finally, if  $\gamma(\mathcal{T}) = 1$ , then any vertex of  $\mathcal{T}$  can follow any other vertex. For  $i \neq j$ , it is clear that the shortest vertex path in  $\mathcal{T}_{[m]}$  from  $(i, \dots, i)$  to  $(j, \dots, j)$  has length  $m + 1$ , so  $\gamma(\mathcal{T}_{[m]}) \geq m = \gamma(\mathcal{T}) - 1 + m$ .  $\square$

**Example 23.** We return to Example 14. The vertices of  $\mathcal{T}_{[2]}$  are  $(1, 1)$ ,  $(1, 2)$ ,  $(2, 3)$ ,  $(2, 4)$ ,  $(3, 1)$ ,  $(3, 2)$ ,  $(4, 3)$ , and  $(4, 4)$ —one for each edge (or vertex path of length 2) in  $\mathcal{T}$ . There are nine edges in  $\mathcal{I}_{[2]}$ . They can be seen in the graph  $\mathcal{T}_{[2]}\mathcal{I}_{[2]}$  in Figure 6.

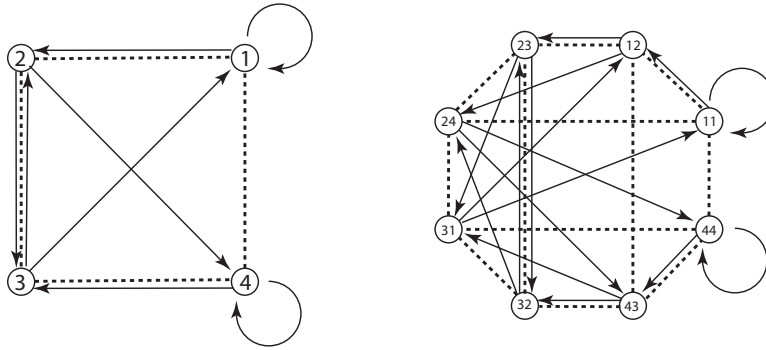


FIGURE 6.

Since  $\gamma(\mathcal{T}) = 2$ , Proposition 22 says that  $\gamma(\mathcal{T}_{[2]}) = 3$ , as we can check by observing that  $A_{[2]}^2$  has zero entries while  $A_{[2]}^3$  does not (where  $A_{[2]}$  is the adjacency matrix for  $\mathcal{T}_{[2]}$ ).

Notice that the set of vertices  $\{(1, 1), (2, 3), (3, 1), (4, 3)\}$  is a maximum cardinality independent set in  $\mathcal{I}_{[2]}$ , so  $\text{ind}(\mathcal{I}_{[2]}) = 4$ . Theorems 3, 9, and 20 tells us that

$$h(f) \geq h_{\mathcal{I}}(\Sigma_{\mathcal{T}}) = h_{\mathcal{I}_{[2]}}(\Sigma_{\mathcal{T}_{[2]}}) \geq \frac{\log 4}{3} = \frac{2 \log 2}{3},$$

which is greater than  $\frac{1}{2} \log 2$ , the lower bound we obtained in Example 14 using  $\mathcal{T}$ . In this particular example we see that  $\gamma(\mathcal{T}_{[m]}) = m + 1$  and  $\text{ind}(\mathcal{I}_{[m]}) = 2^m$ . So

$$h(f) \geq \frac{\log 2^m}{m + 1} = \frac{m \log 2}{m + 1}$$

for all  $m$ . This implies that  $h(f) \geq \log 2$ , the entropy of the doubling map.

## 7. ENTROPY AS A LIMIT

In Example 23 we see that we obtain a very good entropy bound by looking at the sequence of higher shifts. This is true for more than this one example. In this section we show that when the shift space with overlaps is given by a graph  $\mathcal{T}\mathcal{I}$ , the sequence of higher shifts is closely related to the entropy. In particular, if  $\mathcal{T}$  is primitive, the entropy can be expressed as a limit.

As in Section 6, let  $\mathcal{T}_{[m]}\mathcal{I}_{[m]}$  denote the graph for the  $m$ th higher shift. In this case  $\mathcal{B}_m(\Sigma) = V(\mathcal{T}_{[m]}\mathcal{I}_{[m]})$ . Moreover, an  $m$ -separated set  $B$  in  $\Sigma$  corresponds to an independent set in  $\mathcal{I}_{[m]}$ , and hence  $|\mathcal{B}_m^{\text{sep}}(\Sigma, \mathcal{I})| = \text{ind}(\mathcal{I}_{[m]})$ . Thus  $h_{\mathcal{I}}(\Sigma)$  is the growth rate of the independence number of  $\mathcal{I}_{[m]}$  as  $m$  goes to infinity.

**Theorem 24.** *If  $(\Sigma, \mathcal{I})$  is the shift space with overlaps associated to the graph  $\mathcal{T}\mathcal{I}$ , then*

$$h_{\mathcal{I}}(\Sigma) = \limsup_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m}.$$

*If  $\mathcal{T}$  is primitive, then*

$$h_{\mathcal{I}}(\Sigma) = \lim_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})} = \lim_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m}.$$

*(In particular, both of these limits exist.)*

*Proof.* The limit superior equality follows from the definition of entropy and the discussion preceding this theorem. Suppose  $\mathcal{T}$  is primitive with index of primitivity  $\gamma(\mathcal{T})$ . Then  $\mathcal{T}_{[m]}$  is also primitive and  $\gamma(\mathcal{T}_{[m]}) = \gamma(\mathcal{T}) + m - 1$ . So it follows that

$$h_{\mathcal{I}}(\Sigma) = \limsup_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m} = \limsup_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})}$$

Clearly

$$\limsup_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m} \geq \liminf_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m} = \liminf_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})}.$$

So to complete the proof it is enough to show that

$$\liminf_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m} \geq h_{\mathcal{I}}(\Sigma).$$

Since we have  $\text{ind}(\mathcal{I}_{[m]})$   $m$ -separated words and  $\mathcal{T}$  is primitive, we can make from them  $(\text{ind}(\mathcal{I}_{[m]}))^{\lfloor M/\gamma(\mathcal{T}_{[m]}) \rfloor}$   $M$ -separated words by concatenating, as in the proof of Theorem 9. Thus

$$\text{ind}(\mathcal{I}_{[M]}) \geq \text{ind}(\mathcal{I}_{[m]})^{\lfloor M/\gamma(\mathcal{T}_{[m]}) \rfloor},$$

and therefore

$$\liminf_{M \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[M]}))}{M} \geq \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})}.$$

Since this holds for all  $m$ , we have that

$$\liminf_{M \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[M]}))}{M} \geq \limsup_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})} = h_{\mathcal{I}}(\Sigma).$$

□

**Corollary 25.** *If  $\mathcal{T}$  is primitive, then*

$$h_{\mathcal{I}}(\Sigma) = \sup_{m > 0} \left\{ \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})} \right\}.$$

*Proof.* It is obvious that

$$h_{\mathcal{I}}(\Sigma) \leq \sup_{m > 0} \left\{ \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})} \right\}.$$

We must prove the reverse inequality. To do this, observe that since we have  $\text{ind}(\mathcal{I}_{[m]})$   $m$ -separated words and  $\mathcal{T}$  is primitive, we can make them into  $\text{ind}(\mathcal{I}_{[m]})^k$  words of length  $k\gamma(\mathcal{T}_{[m]})$  (for any  $k$ ) by concatenating. Thus  $\text{ind}(\mathcal{I}_{[k\gamma(\mathcal{T}_{[m]})]}) \geq \text{ind}(\mathcal{I}_{[m]})^k$ , so

$$\frac{\log(\text{ind}(\mathcal{I}_{[k\gamma(\mathcal{T}_{[m]})]})}{k\gamma(\mathcal{T}_{[m]})} \geq \frac{\log((\text{ind}(\mathcal{I}_{[m]}))^k)}{k\gamma(\mathcal{T}_{[m]})} = \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})}.$$

By taking the limit of both sides as  $k \rightarrow \infty$ , we obtain

$$h_{\mathcal{I}}(\Sigma) \geq \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})};$$

since this holds for any  $m$ , we are done. □

**Example 26.** Recall from Example 23 that  $\mathcal{T}_{[m]}$  is a primitive graph with index of primitivity  $m + 1$  and independence number  $2^m$ . By Corollary 25,

$$h(f) \geq h_{\mathcal{I}}(\Sigma) = \sup \left\{ \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})} \right\} = \sup \left\{ \frac{\log 2^m}{m+1} \right\} = \sup \left\{ \frac{m \log 2}{m+1} \right\} = \ln 2.$$

Theorem 24 gives us the same bound,

$$h(f) \geq h_{\mathcal{I}}(\Sigma) = \lim_{m \rightarrow \infty} \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{m} = \lim_{m \rightarrow \infty} \frac{\log 2^m}{m} = \ln 2,$$

and happens to be the limit of a constant sequence.

The following example shows that the sequence  $\left\{ \frac{\log(\text{ind}(\mathcal{I}_{[m]}))}{\gamma(\mathcal{T}_{[m]})} \right\}$  may not be nondecreasing.

**Example 27.** Consider the graph  $\mathcal{TI}$  shown in Figure 7.



FIGURE 7.

It is easy to check by hand that

$$\frac{\log(\text{ind}(\mathcal{I}_{[1]}))}{\gamma(\mathcal{T}_{[1]})} = \frac{1}{2} \ln(2) > \frac{1}{3} \ln(2) = \frac{\log(\text{ind}(\mathcal{I}_{[2]}))}{\gamma(\mathcal{T}_{[2]})}.$$

Thus the sequence is not nondecreasing.

We used a computer to find the first several terms of the sequence; based on this, the pattern appears to be  $\left\{ \frac{1}{n+1} \ln(2^m) \right\}_{n=1}^{\infty}$  where  $m = n/2$  for  $n$  even and  $m = (n+1)/2$  for  $n$  odd. Simplified, the  $n$ th term becomes  $\frac{n \ln 2}{2(n+1)}$  when  $n$  is even and  $\frac{1}{2} \ln(2)$  when  $n$  is odd. The two values given by the theorem (the supremum and the limit superior) are equal and the limit superior is equal to the limit. In particular, this common value,  $\frac{1}{2} \ln(2)$  was the first term in the sequence!

## REFERENCES

1. A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*, Classics in Applied Mathematics, vol. 9, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994, Revised reprint of the 1979 original. MR MR1298430 (95e:15013)
2. J. Kwapisz, *Cocyclic subshifts*, Math. Z. **234** (2000), no. 2, 255–290. MR MR1765882 (2001j:37025)
3. ———, *Transfer operator, topological entropy and maximal measure for cocyclic subshifts*, Ergodic Theory Dynam. Systems **24** (2004), no. 4, 1173–1197. MR MR2085908 (2005d:37018)

4. D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995. MR MR1369092 (97a:58050)
5. D. Richeson and J. Wiseman, *Symbolic dynamics for nonhyperbolic systems*, Proc. Amer. Math. Soc. **138** (2010), 4373–4385.
6. C. Robinson, *Dynamical systems*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995. MR 97e:58064
7. A. B. Sossinsky, *Tolerance space theory and some applications*, Acta Appl. Math. **5** (1986), no. 2, 137–167. MR 823824 (87d:92070)
8. A. Szymczak, *The Conley index for decompositions of isolated invariant sets*, Fund. Math. **148** (1995), no. 1, 71–90. MR 96m:58154

DICKINSON COLLEGE, CARLISLE, PA 17013  
*E-mail address:* druckerf@dickinson.edu

DICKINSON COLLEGE, CARLISLE, PA 17013  
*E-mail address:* richesod@dickinson.edu

AGNES SCOTT COLLEGE, DECATUR, GA 30030  
*E-mail address:* jwiseman@agnesscott.edu